

Throughput and Delay Performance Analysis of Packet Aggregation Scheme for PRMA

Qi Zhang Villy B. Iversen
Department of Communication, Optics & Materials
Technical University of Denmark
Lyngby, Denmark
Email: [qz, vbi]@com.dtu.dk

Frank H.P. Fitzek
Department of Communications Technology
Aalborg University
Aalborg, Denmark
Email: ff@es.aau.dk

Abstract—Packet reservation multiple access (PRMA) protocol is an implicit reservation MAC protocol. It is initially designed for voice packets in the cellular networks [2], [3] but it is currently also used for data packets in OFDM based fixed wireless access networks [8], [9]. When it is applied for data packets, the system throughput depends on the size of packets and the number of consecutive packets. From the statistics of existent wireless data networks using PRMA protocol, it shows that the system throughput is quite low because of the inconsecutive small packets. In order to improve the throughput, packet aggregation scheme is considered to be applied in PRMA. Before designing packet aggregation algorithm, it is worth investigating the effect of packet aggregation scheme on the performance of throughput and delay. In this paper we develop a generic Markov chain model for PRMA with packet aggregation. Based on this model the throughput and delay are derived and analyzed. A numerical example is calculated, which illustrates the effect of packet aggregation on the throughput and delay with varying packet arrival rate. The results of the paper are valuable inputs for designing optimal packet aggregation algorithm, considering the tradeoff between throughput and delay.

I. INTRODUCTION

To achieve high throughput values at the network layer, research should focus not only on higher physical layer data rates but also on more efficient media access control (MAC) strategies as the latest WLAN evolution has shown. Besides WLAN many existent fixed wireless access (FWA) networks also have the same problem on the low uplink throughput due to inefficient MAC protocols. Therefore the design of efficient MAC protocols has great significance to improve the overall system throughput at IP level. The advantage of reservation schemes over random access schemes is based on the fact that reservation schemes reduce access collisions and thereby improve the throughput. But reservation schemes require more central control which results in extra signalling overhead and bandwidth scheduling complexity. So far the reservation based MAC protocols can generally be divided into two categories: *explicit* and *implicit* reservation MAC protocols. Reservation ALOHA protocol (R-ALOHA) proposed by Crowther *et al.* in 1973 [1] is a representative explicit reservation MAC protocol. The explicit reservation protocols like R-ALOHA group time into frame and the time slots are divided into explicit contention slots and data slots. Terminals send small bandwidth request packets in the explicit small contention mini

slots. In the explicit reservation MAC protocol, the division of contention slots and data slots duration in the uplink subframe has significant effect on the system throughput and delay performance. Packet reservation multiple access (PRMA) protocol proposed by Goodman *et al.* in [2], [3] is a typical implicit reservation MAC protocol, which is widely used in cellular and satellite networks [6], [5], [4]. In PRMA, there are no explicit contention slots. All the slots in the uplink subframe are of the same type. If a terminal has only one packet to send, PRMA degrades to Slotted-ALOHA (S-ALOHA). If the terminal has more than one packet to send, the terminal sends the first data packet by contention which is implicitly used as reservation request packet. Then the base station broadcasts the outcome of the contention in the next downlink subframe. So the terminals know the time slots usage information (i.e., "reserved" or "available"). The terminals which successfully obtained reservation grant from the base station can transmit packets directly in the reserved slots in the future frames. And the terminals who have packets in buffer but without reservation grant contend the available slots using S-ALOHA. PRMA is a sort of random access protocol but with little central control. Hence PRMA protocol has the advantage of both decentralized packet contention protocols and explicit reservation protocols [10], namely, with little central signalling complexity and good throughput performance. But the throughput performance of PRMA depends on the number of consecutive packets after reservation [4].

Initially PRMA and many PRMA variations [2], [6], [5], [4] were proposed for voice packets in cellular systems. There is a sequence of voice packets with constant size and arrival interval at the terminal after talkspurt starts. Therefore the PRMA always has better throughput performance than S-ALOHA. Currently PRMA is also used for data packets in OFDM based FWA networks due to its simplicity. As data packets have various packet sizes, the big packets are fragmented into several *unit packets*. One unit packet can be transmitted in one slot and so several consecutive uplink subframes complete a packet transmission. So the throughput of the system depends on the size of data packet, namely, the larger packets are the higher throughput is (i.e., the more consecutive unit packets are, the higher throughput is). The worst case is that the

terminals send small packets inconsecutively. In such case the maximum throughput of PRMA is same as that in S-ALOHA only 36.8%. We propose packet aggregation scheme for PRMA to improve throughput. To our knowledge no one has addressed packet aggregation scheme for PRMA before. The main difference of packet aggregation compared with conventional one is that it has more packets to send after it catches the channel. Therefore, it can reduce the average collision probability of each packet and improve the maximum throughput or the channel utilization. But it is also obvious that packet aggregation will eventually result in packet aggregation delay. And to some extent, it also has effect on the terminal channel access delay. So it is very important to investigate the effect of packet aggregation on system throughput and packet delay. Furthermore, it is also worth finding out the threshold of packet aggregation considering the tradeoff between the packet delay constraints and the system throughput.

II. THE PRMA MODEL

In this section we describe a Markov model for PRMA data transmission.

A. Traditional PRMA Model

We assume a PRMA communication system with N homogeneous independent terminals, M slots per frame and p is the permission probability which is the probability that a terminal attempts to transmit a packet in an available slot when a terminal is in state *Cont*. Any terminal can have the following three states:

- Sil the silent state;
- Cont the contending state;
- Tra the transmission state;

The states can change according to the terminal's Markov state transition diagram in Fig. 1. The state transition probabilities are described as following:

- τ one frame period including an uplink subframe and a downlink subframe
- γ the transition probability of a terminal from state *Sil* to state *Cont* at the instant just before uplink subframe starting: If the terminal has no packet in the buffer when uplink subframe starts, it stays in state *Sil*; otherwise, it leaves state *Sil* and enters state *Cont*. We assume the packet arrival process is poisson process with arrival rate λ . So in conventional PRMA, $\gamma = 1 - e^{-\lambda\tau}$.
- σ the probability of a terminal successfully transiting from state *Cont* to state *Tra* during one uplink subframe duration: We assume each terminal only attempts once during one uplink subframe. If a terminal successfully transmits the first packet, it transits from state *Cont* to state *Tra*. Note that because the terminal knows the contention feedback until it receives the feedback contained in the next downlink subframe, this transition happens at the time instant of the downlink subframe ending (i.e, the start of the uplink subframe).

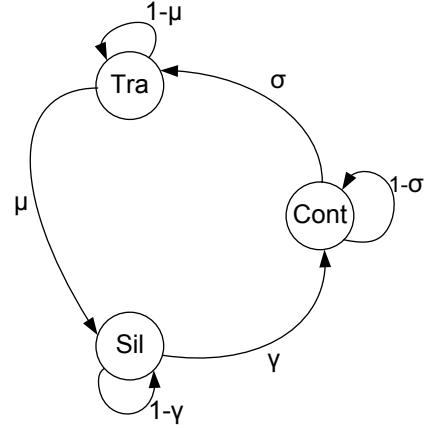


Fig. 1. PRMA terminal state transition diagram

μ the transition probability of a terminal from state *Tra* back to state *Sil* at the instant of the next uplink subframe starting: If there is no packet in its buffer when the next the uplink subframe starts, the terminal lets the reserved slot idle and returns state *Sil*. This is the way how the terminal releases channel reservation in the conventional PRMA protocol. μ depends on the number of consecutive packets in the terminal's buffer, which follows a geometric distribution.

We derive our extended PRMA model based on the general model presented in [10]. We model the PRMA data system as a discrete time Markov process with system state space Ω and one-step transition probability matrix \mathbf{P} .

$$X = \{X_n = (S_n, C_n, T_n) | n \geq 0\} \quad (1)$$

where, S_n , C_n and T_n denote the number of terminals in state *Sil*, *Cont*, *Tra* at time n , respectively. The state transition takes place only at the instant of uplink subframe starting. Due to the the relations among these three variables, the state space Ω is given by

$$\Omega = \{(s, c, t) | s, c, t \geq 0, s \leq N, t \leq \min(M, N), c = N - t - s\} \quad (2)$$

The number of states is $J = (M + 1)(N - M/2 + 1)$, Ω is $J \times 3$ matrix.

The entries of one-step transition probability matrix \mathbf{P} are:

$$P\{X_{n+1} = (s', c', t') | X_n = (s, c, t)\} = \sum_{\substack{s+i-j=s' \\ c+j-k=c' \\ t+k-i=t'}} \alpha_{ijk} \quad (3)$$

where, $\alpha_{ijk} = P_i \cdot P_j \cdot P_k$

- P_i denotes the probability of i transmitting terminals returning to state *Sil* at the transition instant;
- P_j denotes the probability of j silent terminals entering state *Cont* at the transition instant;
- P_k denotes the probability of k terminals contend the channel successfully at the transition instant;

Based on the model of Fig. 1, we can obtain the following expressions for the terms in Eq. 3.

$$P_i = \binom{t}{i} \mu^i (1 - \mu)^{t-i} \quad (4)$$

$$P_j = \binom{s}{j} \gamma^j (1 - \gamma)^{s-j} \quad (5)$$

To obtain α_{ijk} , the most difficult problem is to derive P_k . We use the statistical model based on the model described in [7] to solve it. The model can be summarized as that assuming K contending terminals contending in M slots, how to find the probability of k terminals accessing channel successfully. According to PRMA protocol the probability that a terminal will make a reservation attempt in slot i is denoted by permission probability p_i . Then we assume

$$0 < \sum_{i=1}^M p_i \leq 1 \quad (6)$$

and introduce

$$p_{M+1} = 1 - \sum_{i=1}^M p_i \quad (7)$$

where, p_{M+1} is the probability that the terminal abstains from making a reservation attempt, which we can define as virtual event in the virtual slot (the $M + 1$ th slot). In this paper we assume the terminal are homogeneous so that $p_i = p$ ($1 \leq i \leq M$).

In this model the probability that there exist x_i terminals attempting to transmit packets within slot i ($i=1,2,\dots,M+1$) is given by the Polynomial distribution [7]:

$$P\{x_1, x_2, \dots, x_{M+1}\} = \binom{K}{x_1, x_2, \dots, x_{M+1}} \cdot p_1^{x_1} \cdot p_2^{x_2} \cdot \dots \cdot p_{M+1}^{x_{M+1}}$$

A packet transmission attempt is successful in a slot if one and only one terminal attempts to transmit a packet in this slot. If we can find the distribution of the total number m of successful transmissions ($m=0,1,\dots,M$), we can get the probability of k terminals obtaining the slots and transmitting packets successfully. As it is also very difficult to calculate the distribution of successful transmission directly, We use an effective algorithm was proposed in [7] for resolving this problem. In this algorithm, $P_M(k, x)$ is defined as the probability of k successes out of x attempts within the first available M slots. The detailed derivation of $P_M(k, x)$ can be found in [7]. Hence, when there are $c + j$ terminals in state *Cont*, the aiming probability is

$$P_k = \sum_{x=k}^{c+j} P_M(k, x) \quad (8)$$

The stationary probability distribution

$$\boldsymbol{\pi} = [\pi_{s,c,t}] \quad (s, c, t) \in \boldsymbol{\Omega}$$

can be calculated from transition probability matrix \mathbf{P} . $\boldsymbol{\pi}$ is a vector with J elements.

Let S , C and T denote the number of terminals in each stationary state. So the expectation of S , C and T are listed in the following:

$$\begin{aligned} E[S] &= \sum_{i=1}^J \Omega_{i,1} \cdot \pi_i \\ E[C] &= \sum_{i=1}^J \Omega_{i,2} \cdot \pi_i \\ E[T] &= \sum_{i=1}^J \Omega_{i,3} \cdot \pi_i \end{aligned} \quad (9)$$

where $J = (M + 1)(N - M/2 + 1)$, the number of states. $\Omega_{i,j}$ denote the element at position (i, j) in the state space matrix.

B. Model for PRMA with Packet Aggregation

The idea of packet aggregation strategy is that the terminal does not attempt to transmit packet immediately when a packet arrives. Let n denote the threshold of the number of packets to aggregate. The terminal makes an attempt to transmit when it has at least n packets in its buffer. We use $p_0, p_1, p_2, \dots, p_{n-1}$ to denote the probability that there is 0, 1, 2, \dots , $n-1$ packets in the buffer at the instant of uplink subframe starting. Based on the Fry's equation [11], the expressions and their relations of $p_0, p_1, p_2, \dots, p_{n-1}$ ¹ are given as following. We also assume packet arrival process is a poisson process with arrival density λ . We define γ' as the probability of a single terminal transition probability from state *Sil* to state *Cont*. After we solve $p_0, p_1, p_2, \dots, p_{n-1}$, we can obtain γ' by Eq. 10.

$$\begin{aligned} p_0 &= p_0 p_{\tau,0} + \left(1 - \sum_{i=0}^{n-1} p_i\right) p_{\tau,0} \\ p_1 &= p_0 p_{\tau,1} + p_1 p_{\tau,0} + \left(1 - \sum_{i=0}^{n-1} p_i\right) p_{\tau,1} \\ &\dots \\ p_{n-1} &= p_0 p_{\tau,n-1} + p_1 p_{\tau,n-2} + \dots + p_i p_{\tau,n-1-i} + \dots \\ &\quad + p_{n-1} p_{\tau,0} + \left(1 - \sum_{i=0}^{n-1} p_i\right) p_{\tau,n-1} \\ \gamma' &= 1 - \sum_{i=0}^{n-1} p_i \end{aligned} \quad (10)$$

where, $p_{\tau,j} = \frac{(\lambda\tau)^j}{(j)!} e^{-\lambda\tau}$

The probability P_i^{aggr} and P_k^{aggr} can be calculated in the same way as P_i and P_k . The probability that j silent terminals transit into state *Cont* is:

$$P_j^{aggr} = \binom{s}{j} \gamma'^j (1 - \gamma')^{s-j} \quad (11)$$

When we get the probability P_i^{aggr} , P_j^{aggr} and P_k^{aggr} , the rest calculation of the average number of terminals in the state *Sil*, *Cont*, *Tra* (i.e. $E[S]_{aggr}$, $E[C]_{aggr}$, $E[T]_{aggr}$) can

¹when $n > 6$, the probability of p_6, \dots, p_{n-1} is less than 10^{-4} , therefore, we approximate them as zero to reduce the number of equations and to simplify the calculation.

be calculated by the same approach which is explained in Subsection II-A

III. PERFORMANCE ANALYSIS OF PRMA WITH PACKET AGGREGATION

A. Throughput Analysis

$E[T]_{aggr}$ indicates the average number of slots used in one uplink subframe with packet aggregation. According to $E[T]_{aggr}$, it is easy to get the system throughput or channel utilization, i.e., $E[T]_{aggr}/M$. When the threshold of the number of packets to aggregate is reduced to 1, it becomes conventional PRMA.

B. Delay Analysis

With packet aggregation, the *packet delay* we are interested is the delay experienced by the first arrival packet. So it should be clear that in the following paper, the *packet delay* means the delay experienced by the first arrival packet, i.e., the time from it arriving at a terminal's buffer until the starting instance it can be successfully transmitted. It consists of two parts: aggregation delay and channel access delay.

The aggregate delay is the waiting time of the first arrival packet staying in the buffer, which is equal to the holding time of terminal transiting from state *Sil* to state *Cont*. It is a Erlang-k distribution, so the mean aggregation delay is given by:

$$D_{aggr} = \begin{cases} 0 & n = 1 \\ \tau \left[\frac{\int_0^\infty t \frac{(\lambda t)^{n-1}}{(n-1)!} \lambda e^{-\lambda t} dt}{\tau} \right] = \tau \left[\frac{n-1}{\lambda \tau} \right] & n \geq 2 \end{cases} \quad (12)$$

The channel access delay is the time a terminal stays in state *Cont*. It is calculated based on the model in [10]. Assume the terminal R to be in state *Cont* and S'_n, C'_n and T'_n the number of the rest $N-1$ terminals in the three possible states at time n . We model the PRMA system with the terminal R in state *Cont* by a discrete time Markov chain.

$$X' = \{X'_n = (S'_n, C'_n, T'_n) | n \geq 0\} \quad (13)$$

This Markov chain has state space Ω' and one-step transition probability matrix \mathbf{P}' . The state space Ω' is

$$\Omega' = \{(s', c', t') | s', c', t' \geq 0, s' \leq N-1, t' \leq \min(N-1, M), c' = N-1-t'-s'\} \cup \{\text{Tra}\} \quad (14)$$

where $\{\text{Tra}\}$ is the absorbing state. (s', c', t') is the possible state combination of the $N-1$ terminals when the isolated terminal R is in state *Cont*. So the number of states is $J' = (M+1)(N-M/2)+1$. The one-step transit probability matrix is

$$\mathbf{P}' = \begin{pmatrix} \tilde{\mathbf{P}} & \mathbf{c}_{Tra} \\ 0 & 1 \end{pmatrix} \quad (15)$$

where $\tilde{\mathbf{P}} = [\tilde{P}_{(s'_1, c'_1, t'_1)(s'_2, c'_2, t'_2)}]$, $(s'_1, c'_1, t'_1)(s'_2, c'_2, t'_2) \in \Omega' - \{\text{Tra}\}$. $\tilde{\mathbf{P}}$ is a substochastic matrix, describing the transition probabilities among the possible state combinations of the

TABLE I
PARAMETERS ASSUMPTION IN THE EXAMPLE

Notation	Value
M : slots/ uplink subframe	6
N : number of terminals	12
λ : arrival packets per time τ	[0.02 - 2]
μ : transmission departure probability	$[\leq 1, \leq 0.5, \leq 0.25, \leq 0.167]$
p : permission probability	1/6
τ : frame period	2.304 ms

rest $N-1$ terminals. Column vector $\mathbf{c}_{Tra} = [c_{(s', c', t')Tra}]$, $(s', c', t') \in \Omega'$, is the absorbing probabilities to state *Tra*.

The stationary probability distribution

$$\boldsymbol{\pi}' = [\pi'_{s', c', t'}] \quad (s', c', t') \in \Omega' - \{\text{Tra}\}$$

can be calculated from sub-stochastic matrix $\tilde{\mathbf{P}}$, similarly as in Subsection II-A

This is an absorbing Markov system with one absorbing state. The *fundamental matrix* for the absorbing system is the matrix:

$$\mathbf{Q} = (\mathbf{I} - \tilde{\mathbf{P}})^{-1}$$

The time to absorption in an absorbing Markov system can be obtained from the *fundamental matrix* \mathbf{Q} . The total number of steps expected before absorption equals the total number that it visits all the non-absorbing states. For instance, starting from state i , the total number of step it needs to be absorbed is the sum of all the entries in the i th row of matrix \mathbf{Q} (denoted by q_i). So the channel access delay is equal to the expectation of the time to absorption.

$$D_{access} = \tau \cdot \sum_{i=1}^{J'-1} q_i \cdot \pi'_i$$

So the *packet delay* of the first arrival packet is the the sum of D_{aggr} and D_{access} .

IV. NUMERICAL EXAMPLE

In this section a numerical example is calculated based on the previous described Markov model of PRMA with packet aggregation strategy. The parameters are listed in Table I which are based on the OFDM-based *Motorola Expedience* system. The arrival rate λ means the number of arrival packets during time τ . Since the departure probability μ is assumed as geometric distribution, when μ is in range of $(0, 1.0]$, $(0, 0.5]$, $(0, 0.25]$ or $(0, 0.167]$, respectively, it means that the terminals reserve the same slot at least in 1, 2, 4, 6 consecutive frames.

The system throughput performance is shown in Fig. 2. From Fig. 2 it can be seen that the throughput varies with the packet arrival rate increasing. When arrival rate is low (< 0.2), the achievable throughput has no difference with packet aggregation or not. But the advantage of packet aggregation for improving the throughput can be easily seen, when arrival rate is higher. The different throughput curves reach their peak values at different arrival rate points. For example, the throughput of one packet per contention reaches its maximum value 36.8% at the point of the arrival rate equal to 0.3; the throughput of two packets per contention reaches

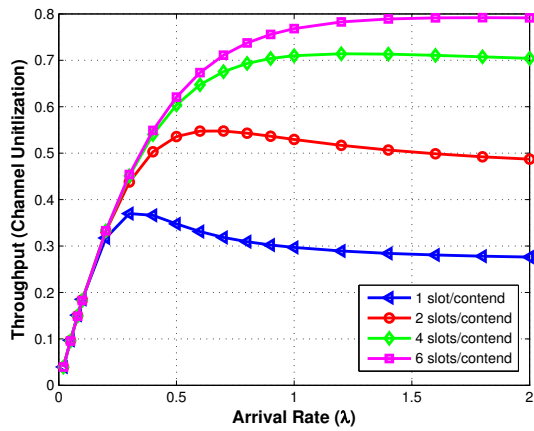


Fig. 2. Packet aggregation effect on system throughput

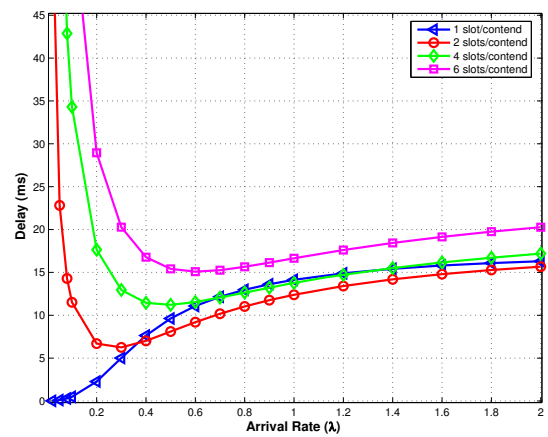


Fig. 3. Packet aggregation effect on the 1st arrival packet's delay

its maximum value 54.7% when the arrival rate is equal to 0.6.

The *packet delay* of the first arrival packet including aggregation delay and channel access delay is shown in shown in Fig. 3. It is clear that it has long packet aggregation delay when the packet arrival rate is very low, namely the first arrival packet should wait longer time in the buffer before the number of packets in the buffer reaches packet aggregation threshold. It is also obvious that with the same arrival rate, it has higher aggregation delay with bigger threshold of packet aggregation. On the other hand, the channel access delay increases with increasing packet arrival rate. Therefore, the *packet delay* decreases first when packet arrival rate increases due to the effect of decreasing aggregation delay. The packet delay will slightly increase because of the effect of the increasing channel access delay when the packet arrival rate getting higher. From Fig. 2 and Fig. 3, they also illustrate that when the packet arrival rate is low, it can not achieve throughput gain by packet aggregation, furthermore it has high delay expense. Therefore it is not worth doing packet aggregation for instance when arrival rate is lower than 0.4 in this example. But when packet arrival rate is high, it can achieve very high throughput by packet aggregation with low *packet delay*. This packet aggregation delay can also be extended to send packets before the first packet becomes too old. So it is a tradeoff between the throughput and the delay. It can guarantee the minimum throughput with fixed number of packets to aggregate. But considering applications' services requirements, the number of aggregated packets should be adaptive according to the delay constraint.

V. CONCLUSION

In this paper we propose packet aggregation scheme in PRMA to improve throughput for data traffic. According to the proposed scheme, we develop a generic Markov chain model for it. Based on this model the throughput and delay performance are derived and analyzed. A numerical example illustrates the efficiency of packet aggregation for improving the throughput. Furthermore, we analyze the effect of packet

aggregation on throughput and delay with varying packet arrival rate. The results of the paper are valuable inputs for designing optimal packet aggregation algorithm, considering the tradeoff of throughput and delay. For system with voice and data joint traffic, the proposed scheme can be extended and give higher priority for voice traffic to preempt access the channel.

REFERENCES

- [1] D. Walden W. R. Crowther, R. Rettberg and F.Heart. "A system for broadcast communication: Reservation aloha", *Proceedings of 6th International Systems and Science Conference*, pages 596–603, Hawaii, Jan 1973.
- [2] K.T. Gayliard D.J. Goodman, P.A. Valenzuela and B.ramamurthi. "Packet reservation multiple access for local wireless communication", *IEEE Transactions on Communications*, 37(8):885–890, August 1989.
- [3] D.J. Goodman S. Nanda and U. Timor. "Performance of prma: A packet voice protocol for cellular systems", *IEEE Transaction on Vehicular Technology*, 40(3):584–598, August 1991.
- [4] G. Bianchi, F. Borgonovo, L. Fratta, L. Musumeci, and M. Zorzi. "C-PRMA: a centralized packet reservation multiple access for local wireless communications", *IEEE Transactions on Vehicular Technology*, 46(2):422–436, 1997.
- [5] Wen, Jyh-Hong and Chen, Yi-Show and Chang, Tain-Sao and Lain, Jenn-Kaie "Theoretical Analysis and Performance of NC-PRMA Protocol for Multichannel Wireless Networks", *Wireless Personal Communications*, 39(2):199–213, 2006.
- [6] F.N. Pavlidou and N. Batsios. "Performance of CDMA/PRMA as an Access Technique for Integrated Services in a UMTS High Altitude Platform System", *Wireless Personal Communications*, 32(3-4): 319–338, 2005
- [7] V.B. Iversen and L. Staalhagen. "Contention analysis of the extended packet reservation multiple access protocol", *ITC Specialist Seminar on Mobile Systems and Mobility*, pages 275–287, Lillehammer, Norway., March 2000.
- [8] Ka-Chung Lai and Xi-Ren Cao. "Analysis of packet reservation multiple access for data transmission", *GLOBECOM'01. IEEE Global Telecommunications Conference (Cat. No.01CH37270)*, pages 711–15 vol.1, 2001.
- [9] X.J. Li and P.H.J. Chong. "Simulation study of some prma-based protocols with channel reservation for data traffic", *2nd International Conference on Mobile Technology, Applications and Systems*, pages 1–5, 2005.
- [10] Francesco Lo Presti and Vincenzo Grassi. "Markov analysis of the prma protocol for local wireless networks", *Wireless Networks*, 4(4):297–306, 1998.
- [11] T.C. Fry. *Probability and its engineering uses*. page 470, New York, 1928.